Grant agreement no: FP7-610603

European Robotic Pedestrian Assistant 2.0 (EUROPA2)

Start of the project: 01.10.2013
Duration: 3 years

DELIVERABLE 4.1

Improved Loop-Closing for Localization and Mapping

Due date: month 12 (September 2014)

Lead contractor organization: KULeuven

Dissemination Level: PUBLIC

1 Introduction

Within the EUROPA2 project, one of the aims is to use already available map information, for example, the geographic information system maps that model the structure of the environment in OpenStreetMap data and other digital map information from commercial providers such as GeoAutomation. A key capability towards this goal is thus high precision localization in publicly available maps. Localization here will also exploit developments done in the context of loop-closing and use them for finding an initial match of the robots observations to the model.

Deliverable 4.1 discusses the first step to improving loop closure. This work is carried out as a close collaboration between KULeuven and GeoAutomation. In the current workflow, the surrounding area is mapped using a combination of structure from motion (SfM) and GPS location. Currently, the data is acquired using a van from GeoAutomation but it can be also acquired from the robot itself, since the required sensors are part of the robot hardware specifications. When GPS reception is low, the positioning of the vehicle can only rely on the SfM algorithm using the visual input. For shorter tracks, the SfM can guarantee sufficient accuracy. However, for larger tracks, and especially when the amount of feature tracks decrease (rural areas), unavoidably the positioning of the van will show a drift.

Loop closure aims at (1) compensating the drift by inserting proper correspondences between difference passes in the bundling scheme to re-evaluate camera positions and van locations, and (2) identifying possible correspondences between different passes in an automated way. The sequel of the deliverable addresses the experiments that have been carried out on both of these aspects.

2 Loop bundling

Considering the GeoAutomation processing pipeline, when GPS reception and SfM features extraction are reliable, they jointly guarantee an accurate mapping of the environment executing the processing pipeline in a purely sequential order. When reliability decreases, and drift occurs, the measurements carried out in the different passes on the same area (e.g. a crossing) will differ and lead to inconsistent mapping of the topography. When we create an orthogonal mapping of all the tracks (see Figure 1 for an example) such inconsistencies can be easily visually identified. The inconsistencies can be taking care of by creating links, i.e. point correspondences, between different parts of the tracks that pass on the same spots. In a first stage we assume these correspondences can be indicated manually to get a feeling for the behavior of the whole system when trying to correct for the inconsistencies on the crossings. There is also a practical reason for this; in real situations the landscape is not always very distinctive. For rural environments for instance, such as the one depicted in Figure 2, many crossings look very similar, which severely challenges possible automated correspondence search, as we will discussed in the next section.

Doing so, the camera poses and tracks can be recalculated in a *re-bundling* step. However, considering local maps can be quite complicated and consist of many crossings, introducing all possible links in one single global bundling step substantially slows down the optimization process. Experiments show that the re-bundling is much more efficient when inconsistencies are threated individually, and secondly by limiting the area around the crossing in which the re-bundling is to be executed. Furthermore, convergence is faster when the observations from one pass are (temporarily) fixed while others are being re-iterated.

A large amount of areas have been processed in this manner. Figure 2 shows a selection of

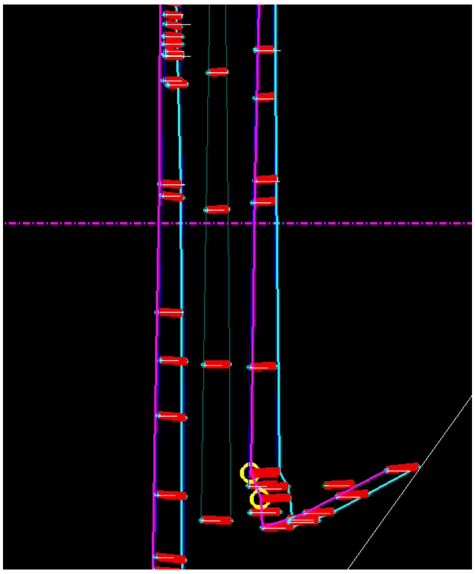

Figure 1: Mapping inconsistencies as seen in the orthographic topographical map, purple and blue lines should match, but have been measured on different passes in the bundling scheme

crossings where measurements have been carried out in one pass, and projected onto the imagery in another pass. The lighter and darker blue and purple lines indicate the borders and delineate roads, traffic markings, pedestrian areas, etc. When drift has occurred, one can observe that these markings do not match the imagery from the other passes. The short red marker lines indicate how these measurements will be shifted, i.e., corrected, in the image plane after re-bundling. One can verify that repositioning of the measurements fits the imagery fairly well. The convergence and final output depends on several factors, the depth accuracy of the corresponding points, the variability in depth information in the overlapping sections, the amount of passes in these areas, etc. Further investigation, in conjunction with the automated loop detection, will be carried out in the next period.

3 Automated loop detection and correspondences

When localizing or building maps of the environment, it is important that revisiting the same spot is noticed by the robot autonomously and that corresponding points in the initial and later 3D reconstructions get precisely matched. At the same time, the capabilities to recognize the same spot should be robust against changes in appearance as can be caused by difference in weather or time of day.

In this section we describe the first steps to match and search for 3D structures that have been recorded at different times. We start from the city models that have been recorded by GeoAutomation in the context of D1.4. As an initial test we select urban models above rural environments because of the possible ambiguity in the latter, as noted in the previous section. Furthermore, the robot will be navigating in an urban environment.

In the city models several regions can be identified where the tracks cross each other (crossings) or run parallel (same road passed in two directions) or both (complexer traffic situations, roundabouts..). The availability of image information allows for additional techniques to be used for accurate mapping. Loop closure, where the system can automatically recognize an area that it has already traveled

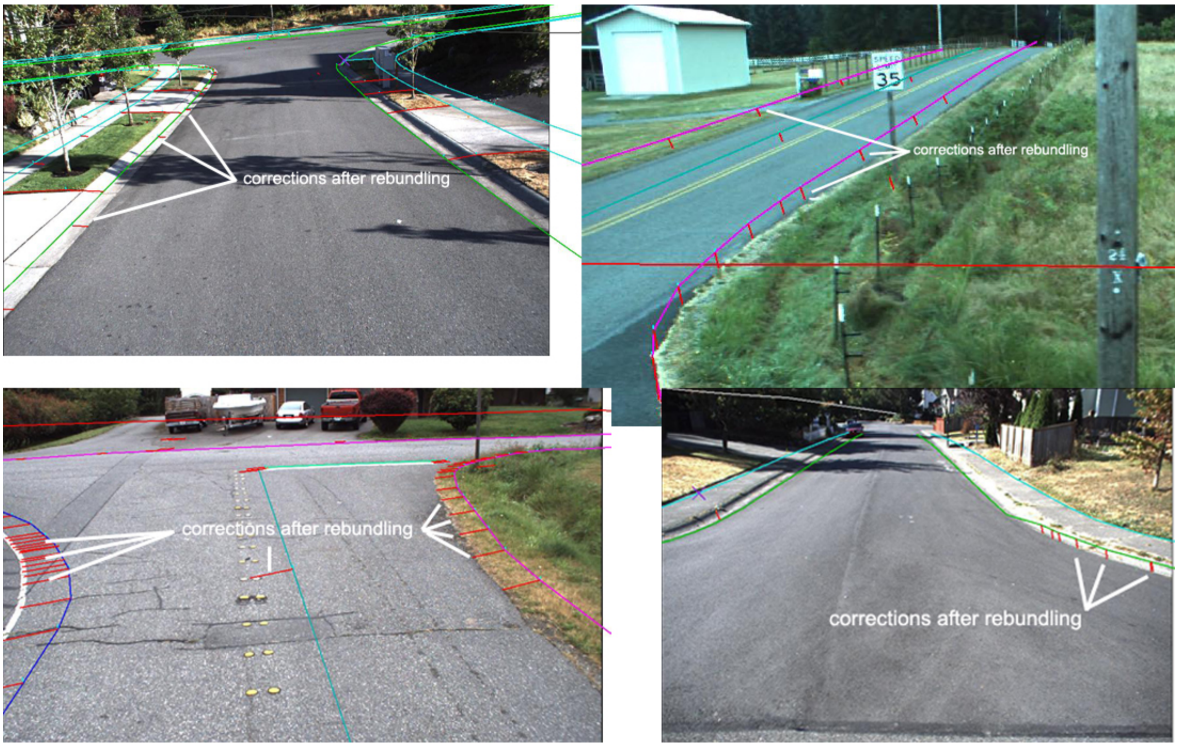

Figure 2: A selection of different crossings from a GeoAutomation dataset (Snohomish), the drift of the original measurements is clearly visible. The red correction lines indicate how these measurements will be repositioned after re-bundling.

Figure 3: The same crossing in two passes, pass1 is seen through the frontal camera, pass2 is seen through a side camera viewpoint.

Figure 4: The Leuven area focussing on two crossings.

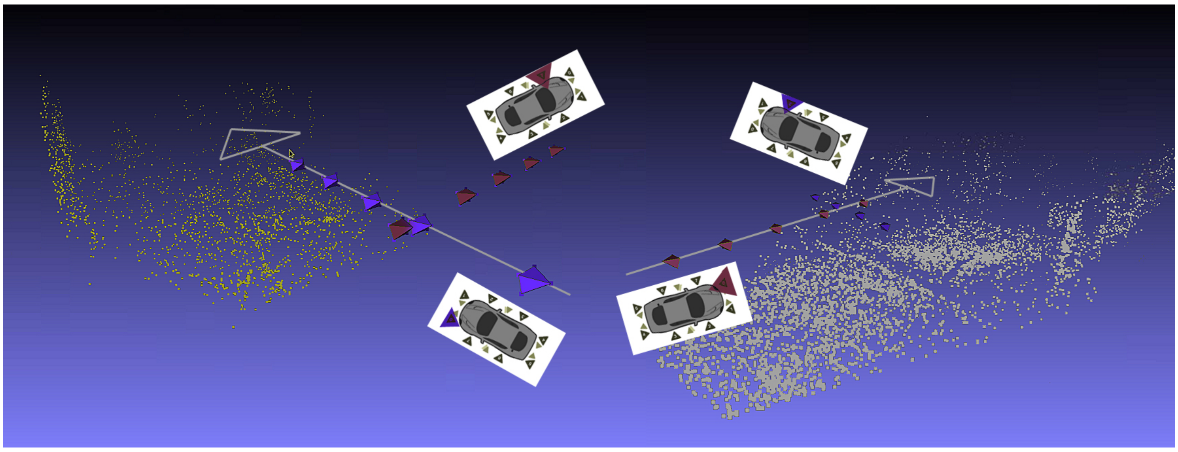

Figure 5: Correspondence and viewpoint alignment for crossing A. Matches are found between the frontal and side camera viewpoints respectively.

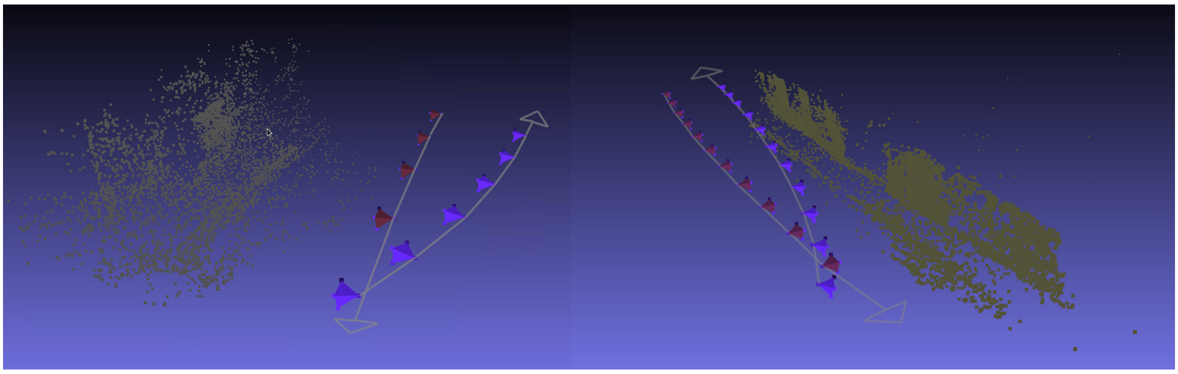

Figure 6: Correspondence and viewpoint alignment for crossing B. Matches are found between the opposite side camera viewpoints respectively.

in the past, can benefit from visual information and used to reduce or even eliminate drift. Over the last decade, important progress has been made, in that features have been devised that can be recognized again even under wide-baseline conditions. Yet, so far only two types of invariance have been considered in order to arrive at such capabilities. There are local features that are invariant under similarity transformations, like SIFT or SURF, and affine invariant features have been derived as well, e.g., MSER. We have used an own implementation of the SURF features in the experiments.

Given the urban context of the dataset, we verified to what extent the current feature matching approach can match amongst different views, even when viewing and lighting parameters are different. For the Leuven dataset, a lot of lighting differences can be observed in the sequence due to bright sun and corresponding shadows. The auto-shutter of the cameras also causes the overall brightness of the image to change over the sequence. Figure 3 gives an impression of two passes through the same crossing, seen through two different camera viewpoints. As another example, Figure 4 focuses on two crossings. In crossing A, the tracks passed each other perpendicularly, in crossing B, the tracks passed each other parallel. Figures 5 and 6 show the results. The corresponding features that are found between the different tracks do not only give an indication that this location has been passed twice, the correspondences also allow to determine the 3D locations of these features, as well as the relative camera positions of both two tracks with respect to one another. The camera viewpoints are indicated with small pyramids. The blue ones indicate the relative positions of the first pass, the red ones for the second pass.

The search is exhaustive still, more research is needed to find proper compact descriptions, generate interest points and their neighborhoods, as well as the descriptions and detectors therein that are covariant and invariant with the affine subgroups.